

# Defection and extortion as unexpected catalysts of unconditional cooperation in structured populations

Attila Szolnoki<sup>1,\*</sup> and Matjaž Perc<sup>2,3,†</sup>

<sup>1</sup>*Institute of Technical Physics and Materials Science, Research Centre for Natural Sciences,  
Hungarian Academy of Sciences, P.O. Box 49, H-1525 Budapest, Hungary*

<sup>2</sup>*Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, SI-2000 Maribor, Slovenia*

<sup>3</sup>*CAMTP – Center for Applied Mathematics and Theoretical Physics,  
University of Maribor, Krekova 2, SI-2000 Maribor, Slovenia*

We study the evolution of cooperation in the spatial prisoner's dilemma game, where besides unconditional cooperation and defection, tit-for-tat, win-stay-lose-shift and extortion are the five competing strategies. While pairwise imitation fails to sustain unconditional cooperation and extortion regardless of game parametrization, myopic updating gives rise to the coexistence of all five strategies if the temptation to defect is sufficiently large or if the degree distribution of the interaction network is heterogeneous. This counterintuitive evolutionary outcome emerges as a result of an unexpected chain of strategy invasions. Firstly, defectors emerge and coarsen spontaneously among players adopting win-stay-lose-shift. Secondly, extortioners and players adopting tit-for-tat emerge and spread via neutral drift among the emerged defectors. And lastly, among the extortioners, cooperators become viable too. These recurrent evolutionary invasions yield a five-strategy phase that is stable irrespective of the system size and the structure of the interaction network, and they reveal the most unexpected mechanism that stabilizes extortion and cooperation in an evolutionary setting.

Widespread cooperation in nature is one of the most important challenges to Darwin's theory of evolution and natural selection, but it is also the main driving force behind the evolutionary transitions that led from single-cell organisms to complex animal and human societies [1]. And it appears to be this mixture of a fascinating riddle and utmost importance that makes cooperation so irresistibly attractive to study. Evolutionary game theory [2–6] is thereby the most frequently employed theoretical framework, revealing mechanisms such as kin selection [7], network reciprocity [8], direct and indirect reciprocity [9, 10], as well as group selection [11] as potent promoters of cooperative behavior. Adding to these established five rules for the evolution of cooperation [12], recent years have witnessed a surge of predominantly interdisciplinary studies, linking together knowledge from biology, sociology, economics as well as mathematics and physics, to identify new ways by means of which the successful evolution of cooperation amongst selfish and unrelated individuals can be understood [13–20].

From the large array of games that make up evolutionary game theory, none has received as much attention as the prisoner's dilemma game [21–35]. Each instance of the game is contested by two players who have to decide simultaneously whether they want to cooperate or defect. The dilemma is given by the fact that although mutual cooperation yields the highest collective payoff, a defector will do better if the opponent decides to cooperate. The rational outcome is thus mutual defection. The popularity of the game was helped significantly by the tournaments that were organized by Robert Axelrod [36], where the most successful strategy for the iterated prisoner's dilemma game was sought. Interestingly the long-term winner was the tit-for-tat strategy by the simple and in-

tuitive virtue of always following the opponent's previous action. However, tit-for-tat cannot correct erroneous moves, and it is also vulnerable to random drift when mutant strategies appear which always cooperate [37]. Nowak and Sigmund therefore proposed win-stay-lose-shift as another equally simple strategy that has neither of these two disadvantages, and can outperform tit-for-tat in the prisoner's dilemma game [22]. Players adopting win-stay-lose-shift simply repeat the previous move if the resulting payoff has met their aspiration level and change otherwise.

The simplicity and effectiveness of strategies like tit-for-tat and win-stay-lose-shift were unrivaled for decades, and they generated a large following of the seminal works that introduced them. Recently, however, Press and Dyson have introduced a new class of so-called zero-determinant strategies that can dominate any opponent in the iterated prisoner's dilemma game [38]. A particularly interesting subset of the class are extortion strategies, which ensure that an increase in one's own payoff exceeds the increase in the other player's payoff by a fixed percentage. Extortion is therefore able to dominate any opponent [39]. But this holds only if players are unable to change strategies in response to their failures. In an evolutionary setting, where players are able to imitate strategies that are more successful, extortion was shown to be evolutionary unstable [40]. If the two players engaged in the game belong to distinct populations, or if the population size is very small, on the other hand, extortioners can nevertheless prevail, and rather counterintuitively, they may also act as catalysts for the evolution of cooperation [41]. Evolutionary stability can also be warranted by generous zero-determinant strategies through their mutually supporting behavior [42].

Results summarized thus far concerning zero-determinant strategies were obtained in well-mixed populations. Yet it is well-known that stable solutions in structured population can differ significantly from those in well-mixed populations. The most prominent example of this fact is the successful evolution of cooperation in the spatial prisoner's dilemma game

\*Electronic address: szolnoki.attila@ttk.mta.hu

†Electronic address: matjaz.perc@uni-mb.si

through network reciprocity [8]. Further examples include the stabilization of reward [43], peer and pool punishment [44, 45], in-group favoritism [46], as well as homophily [47], to name but a few. Indeed, the fact that the interactions among players are frequently not random and best described by a well-mixed model, but rather that they are limited to a set of other players in the population and as such are best described by a network, has far-reaching consequences for the outcome of evolutionary processes [13, 15, 16, 18, 19].

Motivated by this, we have recently shown that in structured populations the microscopic dynamic that governs strategy updating plays a decisive role for the fate of extortioners [48]. By using the simplest three-strategy model, comprising cooperators ( $C$ ), defectors ( $D$ ), and extortioners ( $E_\chi$ ), we have shown that pairwise imitation and birth-death dynamics return the same evolutionary outcomes as reported previously in well-mixed populations. The usage of myopic best response strategy updating, on the other hand, renders extortion evolutionary stable via neutral drift. Counterintuitively, the stability of extortioners helps cooperators to survive even under the most testing conditions, whereby the neutral drift of  $E_\chi$  players serves as the entry point, akin to a Trojan horse, for cooperation to grab a hold among defectors. Although the mutually rewarding checkerboard-like coexistence of cooperators and extortioners can always be temporarily disturbed by defectors, it is only a matter of time before the neutral drift reintroduce extortioners and the whole cycle starts anew.

Here we extend our study to five competing strategies, taking into account also the tit-for-tat strategy ( $TFT$ ) and the win-stay-lose-shift strategy ( $WSLS$ ), in addition to the previous three that we have studied in [48]. The five strategies  $D$ ,  $C$ ,  $E_\chi$ ,  $TFT$ , and  $WSLS$  are the same as studied recently by Hilbe et al. [41] in well-mixed populations, with the strength of the social dilemma  $b$  and the strength of exploitation  $\chi$  being the two main parameters that determine the payoffs amongst the strategies. For details about the parametrization of the game and the applied updating rules, we refer to the Methods section. The inclusion of the tit-for-tat strategy and the win-stay-lose-shift strategy promises fascinating evolutionary outcomes, especially since under well-mixed conditions  $D$  can beat  $WSLS$ , but the dominance reverses in the presence of the other three strategies. As we will show in the next Section, in structured populations  $WSLS$  dominate completely for sufficiently small values of  $b$  if the interaction network is characterized by a homogeneous degree distribution. Beyond a threshold value of  $b$ , or if the interaction network is characterized by a heterogeneous degree distribution (see for example [49]), however,  $D$  emerge and coarsen spontaneously, which in turn opens up the possibility for all the other strategies to emerge as well.

## Results

Before turning to the main results obtained with myopic best response updating, we present in Fig. 1 the evolutionary outcomes obtained via imitation on a square lattice. If imitation is the basis of strategy updating, then neither cooper-

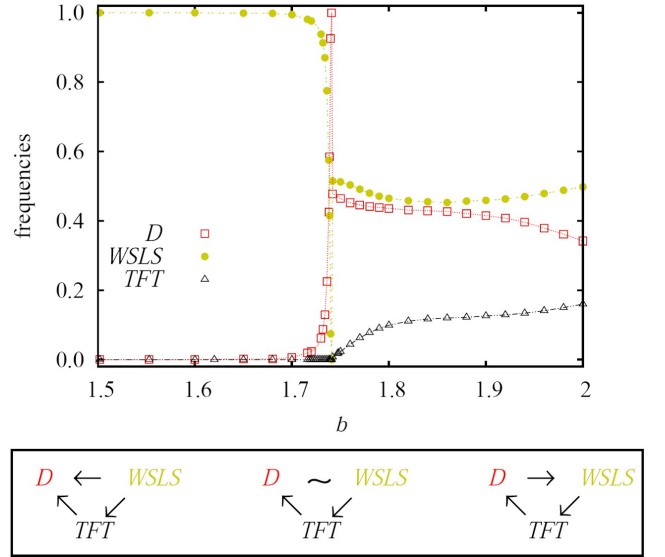


FIG. 1: Imitation on a square lattice fails to sustain cooperation and extortion. Depicted are the stationary frequencies of surviving strategies in dependence on the strength of the social dilemma  $b$ . It can be observed that for sufficiently small values of  $b$  only  $WSLS$  survive. As  $b$  increases, the pure  $WSLS$  phase first gives way to a narrow two-strategy  $WSLS + D$  phase, which then transforms into the three-strategy  $WSLS + TFT + D$  phase. The emergence of these three different phases is a direct consequence of dominance relations between the three involved strategies, which are schematically depicted in the bottom frame for the respective values of  $b$  from left to right. Arrows show the direction of invasion between strategies.

ators nor extortioners can survive, and this regardless of the strength of the social dilemma and the strength of exploitation. Since extortioners always die out, the composition of the final state is actually completely independent of  $\chi$ . We have used  $\chi = 1.5$  for the presented results, but the value influences only the time needed for relaxation towards the final stable solution. Starting with  $b \geq 1$  (we show results from  $b = 1.5$  onwards for clarity with regards to the subsequent phase transitions), the completely dominant strategy is  $WSLS$ . At the other end of the interval of  $b$ , we have a stable three-strategy  $WSLS + TFT + D$  phase, which is sustained by cyclic dominance. In between, we have a narrow two-strategy  $WSLS + D$  phase, which terminates immediately after  $D$  reach dominance.

This dependence on  $b$  can be understood by considering the relations among the surviving strategies, as summarized in the bottom frame of Fig. 1. For small values of  $b$  (left),  $WSLS$  dominate both  $D$  and  $TFT$ . The latter also dominate  $D$ , but their superior status in this relationship has no effect on the final state. For high values of  $b$  (right), the direction of invasion between  $WSLS$  and  $D$  changes compared to the low  $b$  case, while the other two relations remain unchanged. Consequently, instead of a pure  $WSLS$  phase, we have a three-strategy  $WSLS + TFT + D$  phase, where  $WSLS$  invade  $TFT$ ,  $TFT$  invade  $D$ , and  $D$  invade  $WSLS$  to close the loop of dominance. It is worth emphasizing that this solution is im-

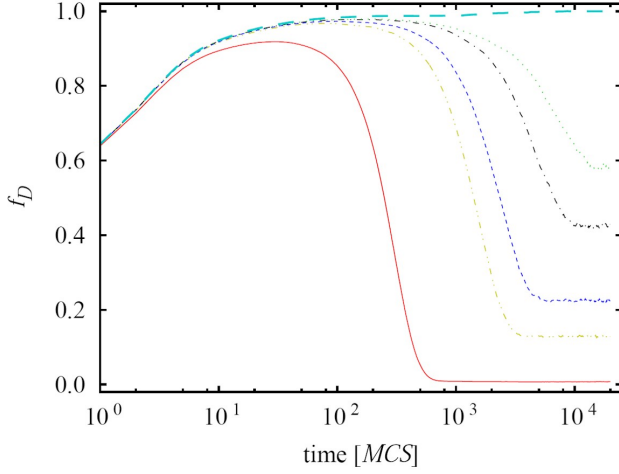


FIG. 2: The coexistence of defectors and players adopting the win-stay-lose-shift strategy in case of imitation on a square lattice. Depicted is the time evolution of the frequency of defectors  $f_D$  as obtained for  $b = 1.7, 1.734, 1.736, 1.738, 1.739$  and  $1.741$  from bottom to top. The time courses provide insight into the competition for space within the narrow two-strategy  $WSLS + D$  phase that can be observed in Fig. 1. At  $b = 1.741$  defectors come to dominate the whole population, but their dominance is immediately overthrown in favor of the three-strategy  $WSLS + TFT + D$  phase that is sustained by cyclic dominance. The used linear size of the square lattice is  $L = 1000$ . Note that the time scale is logarithmic.

possible in a well-mixed population for all  $b < 2$ .

In a narrow interval between the pure  $WSLS$  phase and the cyclic  $WSLS + TFT + D$  phase, we have the situation depicted in the middle of the bottom frame of Fig. 1, where unlike for small and high values of  $b$ , the relation between  $WSLS$  and  $D$  enables their coexistence in a structured population. As for small values of  $b$ , here too  $TFT$  can invade  $D$ , but this is without effect on the final outcome. The stable two-strategy coexistence is illustrated in Fig. 2, where we show how  $WSLS$  and  $D$  compete for space over time for different values of  $b$ . The larger the value of  $b$ , the smaller the fraction of the population that is occupied by  $WSLS$  in the stationary state. Interestingly, when  $b$  is large enough for  $D$  to fully eliminate  $WSLS$ , the complete dominance of defectors is prevented by the presence of  $TFT$ , who become viable via a second-order continuous phase transition. From this point onwards, the cyclic dominance  $WSLS \rightarrow TFT \rightarrow D \rightarrow WSLS$  starts working until the end of the interval of  $b$ , as depicted in the main panel of Fig. 1.

Overall, extortion is unable to capitalize on structured interactions if the strategy updating is governed by imitation or a birth-death rule (results not shown), and in fact this is in full qualitative agreement with the results obtained in well-mixed populations [40, 41]. In the realm of evolutionary games, extortioners do not do well against cooperative strategies like  $C$ ,  $TFT$  and  $WSLS$ . They may thrive for a short period of time, but as soon extortion becomes widespread, it is more profitable to cooperate, which ultimately renders extortion evolutionary unstable.

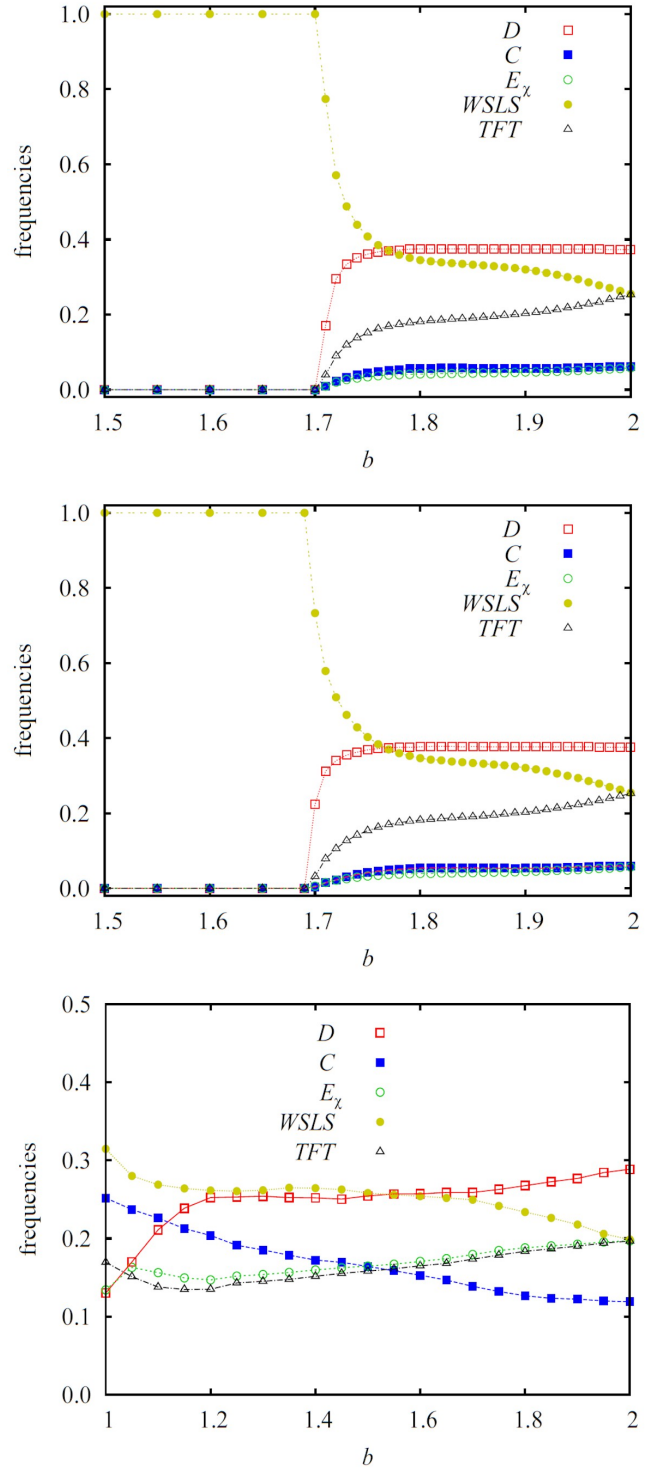


FIG. 3: Myopic best response updating in structured populations stabilizes extortion and cooperation. Depicted are the stationary frequencies of surviving strategies in dependence on the strength of the social dilemma  $b$ , as obtained for the strength of extortion  $\chi = 1.5$  on the square lattice (top), the random regular graph (middle), and the scale-free network (bottom). It can be observed that players adopting the  $WSLS$  strategy dominate for sufficiently small values of  $b$  on homogeneous interaction networks (top and middle), but as  $b$  increases or if the interaction network is heterogeneous (bottom), the pure  $WSLS$  phase gives way to a stable five-strategy  $WSLS + D + E_\chi + TFT + C$  phase. (continues on next page)

FIG. 3: (*continues from previous page*) Here defectors emerge and coarsen spontaneously because for sufficiently large values of  $b$  their payoff becomes larger than that of clustered *WSLS* players. The emergence of defectors immediately opens the door to the survival of extortioners and *TFT* players, which both emerge by chance and spread by means of neutral drift. Lastly, with the emergence of extortioners and *TFT* players cooperators become viable as well, thus forming the stable five-strategy phase. The latter is virtually unaffected by different values of  $\chi$ , as demonstrated in Fig. 4. Importantly, the described coexistence of the competing strategies is a universal behavior that can be observed in structured populations regardless of the properties of the interaction network, and even across the whole span of  $b$  values, as illustrated in the bottom panel. Characteristic snapshots depicting the described key stages of the evolutionary process are presented in Fig. 5.

Myopic strategy updating, on the other hand, can sustain very different evolutionary outcomes as it allows players to adopt strategies that are not necessarily present in their interaction neighborhood. In fact, strategies need not be present in the population at all, as long as they are an option for the players to choose randomly when it is their turn to perhaps change their strategy. Nevertheless, we emphasize that myopic best response updating is different from mutation, because each individual strategy change is still driven by the payoff difference, as described by Eq. 1. Results presented in Fig. 3 obtained on the square lattice (top) and the random regular graph (middle) show that for sufficiently small values of  $b$  the final state is the same as under imitation dynamics. Players adopting *WSLS* dominate completely from  $b = 1$  onwards (as in Fig. 1, we show results for  $b \geq 1.5$  only). At a critical value of  $b$ , however, a second-order continuous phase transition rather unexpectedly leads to the stable coexistence of all five competing strategies. A similar diversity of strategies prevails on heterogeneous interaction networks, as illustrated by the results obtained on a scale-free network shown in the bottom panel of Fig. 3. Myopic best response updating is thus able to stabilize extortion in structured populations. Perhaps even more surprisingly, as the strength of the social dilemma increases, the two cooperative strategies *C* and *TFT* become viable as well. This outcome is rather independent of the structure of the interaction network.

Since extortioners survive for sufficiently high values of  $b$ , the strength of extortion  $\chi$  might play a role too, but as evidenced by the results presented in Fig. 4, this role is in fact very minor. As the value of  $\chi$  increases, the extortioners become slightly more common on the expense of *TFT* and *C* players, but overall this does not affect the evolutionary stability of extortion and cooperation. Compared to our previous results presented in [48], where we have studied the three strategy variant of the game without *TFT* and *WSLS* players, the role of  $\chi$  is less significant here mainly because the stationary frequency of extortioners is much smaller. The fact that their frequency is much smaller, however, is a direct consequence of the presence of the two additional cooperative strategies (*TFT* and *WSLS*), which in turn highlights the general subordinate role of extortioners compared to cooperation in evolutionary games. The latter was emphasized already

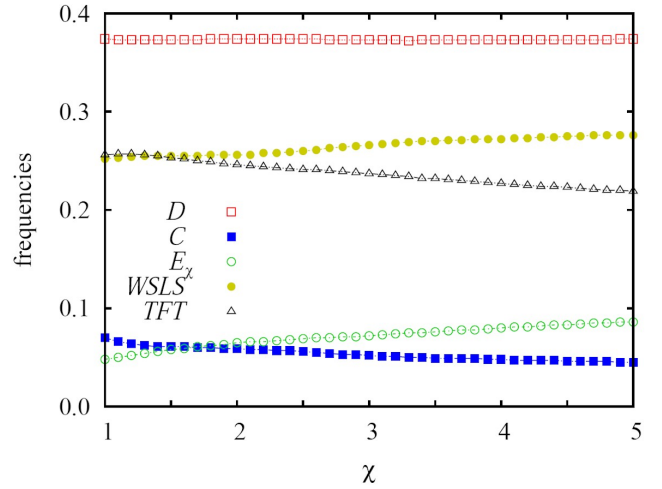


FIG. 4: The strength of extortion has a negligible impact on the stationary frequencies of competing strategies, and it does not affect the evolutionary stability of extortion and cooperation. Depicted are the stationary frequencies of surviving strategies in dependence on the strength of extortion  $\chi$ , as obtained for the social dilemma strength  $b = 2$  on a square lattice. It can be observed that the variations of all frequencies are small. Expectedly, larger values of  $\chi$  favor extortion. The neutral drift of *TFT* players therefore becomes slightly less prolific, which in turn also slightly decreases the frequency of cooperators. Interestingly, the stationary frequencies of strategies at  $b = 2$  and their  $\chi$ -dependency are practically indistinguishable for the square lattice and the random regular graph. This further highlights the irrelevance of the structure of the interaction network under myopic best response updating, and thus also the universality of the presented results.

in [40, 41], as well as by the results presented in Fig. 1 above. Also contributing to the minor role of  $\chi$  is that the emergence of extortioners is in fact a second-order effect, as we will explain next.

To understand why  $E_\chi$ , *TFT* and *C* emerge as  $b$  increases, it is instructive to consider the erosion of the pure *WSLS* phase on square lattice, as illustrated in Fig. 5. For a sufficiently high value of  $b$  defectors emerge and start coarsen spontaneously because their payoff becomes competitive with the payoff of aggregated *WSLS* players. The emergence of the *D* phase, however, paves the way for the emergence of all the other strategies. Namely, both  $E_\chi$  and *TFT* are neutral against *D*, and thus they may emerge by chance and spread via neutral drift. As  $E_\chi$  accumulate locally, *C* become viable too because their payoff is higher. The emergence of *C* is helped further (or at least not hindered) by *TFT*, who are neutral with *C*. During this unexpected chain of strategy invasions, defection and extortion thus emerge as catalysts of unconditional cooperation. Effectively, the defectors act as a Trojan horse for all the other strategies, while subsequently the extortioners act as a Trojan horse for cooperation. Evidently, the spreading of *C*, which utilizes the neutral drift of  $E_\chi$ , will be controlled by defectors and *WSLS* players who can strike back since their presence in place of an extortioner may yield a higher payoff in a predominantly cooperative neighborhood.



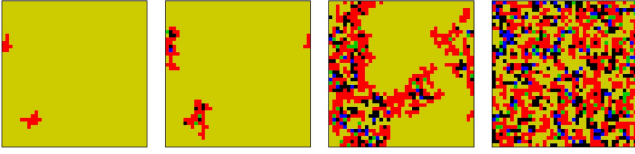


FIG. 5: Characteristic time evolution of the spatial distribution of the five competing strategies on a square lattice. The evolution starts from the full *WSLS* phase (not shown), using  $b = 1.8$  and  $\chi = 1.5$ . At  $MCS = 5$  (leftmost panel), first defectors start emerging because their payoff is comparable with *WSLS* players. Soon thereafter, at  $MCS = 10$  (second panel from left), first extortioners and *TFT* players emerge. Both have neutral relations with the defectors, and thus their emergence and spreading are due to chance and neutral drift. At  $MCS = 30$ , as soon as locally the number of extortioners becomes sufficiently large, cooperators emerge as well due to their higher payoffs, and their spreading is additionally supported by the *TFT* players. The recurrence of these elementary processes eventually spreads a stable mixture of all five strategies across the whole population, as depicted in the rightmost panel that was taken at  $MCS = 100$ . The color encoding of the strategies is the same as used in Figs. 3 and 4. For clarity with regards to individual players and their strategies, we have used a small square lattice with linear size  $L = 40$ .

This, however, will again be only temporary, since the described elementary invasions are bound to recur, thus assuring the stability of the  $WSLS + D + E_\chi + TFT + C$  phase.

An important lesson learned from the presented results in Fig. 5 is that although extortion can be as counterproductive as defection, it is still less destructive. For an unconditional cooperator it never pays sticking with the strategy if surrounded by defectors, but it may be the best option among extortioners. Cooperators are of course happiest among other cooperators, but in the presence of extortioners they can still attain a positive payoff, and this is much better than nothing or a negative value in the presence of defectors. It is worth emphasizing that this argument is valid independently of the properties of the interaction network, as the described chain of strategy invasions emerges in all the structured populations that we have considered.

### Discussion

We have shown that even if the set of competing strategies is extended to encompass, besides unconditional cooperators, defectors and extortioners [48], also the tit-for-tat strategy and the win-stay-lose-shift strategy, the imitation dynamics in structured populations is still unable to render extortion evolutionary stable. For sufficiently small values of  $b$  only players adopting the win-stay-lose-shift strategy survive, while beyond a threshold value a stable three-strategy phase consisting of defectors, tit-for-tat and win-stay-lose-shift players emerges. Since extortioners never survive, the strength of exploitation  $\chi$  is without effect. These results agree with those reported previously for sizable isolated well-mixed populations [41], and they highlight the severe challenges that extortioners face when vying for survival in the realm of evolu-

tionary games where players are able to imitate strategies that are performing better [40].

If the evolution is governed by myopic best response updating, however, the outcomes are significantly different from those obtained via imitation. We have shown that for sufficiently large values of  $b$  the complete dominance of win-stay-lose-shift players is broken as soon as defectors emerge and start coarsening. Subsequently, within the homogeneous domains of defectors, extortion becomes viable too via the same mechanism as we have described before in [48]. In particular, extortioners and defectors are neutral, and hence the former can emerge by chance and spread via neutral drift. Yet as soon as extortioners emerge, cooperators can finally emerge as well, because in competition with the former they are superior. In this evolutionary scenario, defection and extortion thus act as the most surprising catalysts of unconditional cooperation in structured populations. Moreover, we have shown that the coexistence of all competing strategies occurs across the whole interval of  $b$  values if a heterogeneous (scale-free) network describes the interactions among players. Because of this unlikely path towards cooperation, we conclude that defectors and extortioners effectively play the role of a Trojan horse for cooperators. Interestingly, similar transient roles of extortionate behavior were recently reported in the realm of well-mixed populations when studying the adaptive dynamics of extortion and compliance [50]. Moreover, after the emergence and coarsening of defectors, in the presently studied game the tit-for-tat players also become viable as they are likewise neutral, and can thus spread via neutral drift just like extortioners. In recurrence, these evolutionary processes give rise to a stable five-strategy phase that is hardly affected by the strength of exploitation  $\chi$ , and it is also robust to the population size and the structure of the interaction network.

Taken together, these results thus have a high degree of universality and highlight the relevance of coarsening, the emergence of role-separating strategy distributions (which manifests as checkerboard ordering on regular graphs), and best response updating in evolutionary games. The latter is especially important, as it appears to be an integral part of human behavior [51–53]. From the more pragmatical point of view, best response updating conveys to the players an ability to explore the space of available strategies even if they are not present in their immediate neighborhood or even in the population as a whole, and by doing so, such updating dynamics opens up the door to the most counterintuitive evolutionary outcomes. Similarly to kin competition, the presented results also highlight the other side of network reciprocity. Namely, it does not only support cooperative behavior by means of clustering, but it also reveals the consequences of bad decisions – defectors and extortioners become weak when they become surrounded by their like. From this point of view, it is understandable and indeed expected that structured populations, if anything, hinder the successful evolution of extortion under imitation. The surprising positive role of extortioners becomes apparent only under best response updating, where the threatening loom of widespread defection is drifted away by the lesser evil to eventually introduce more constructive cooperative strategies.

## Methods

We adopt the same game parametrization as Hilbe et al. [41]. Accordingly, the payoff matrix for the five competing strategies is

	<i>TFT</i>	<i>WSLS</i>	$E_\chi$	<i>all C</i>	<i>all D</i>
<i>TFT</i>	$\frac{1}{2}$	$\frac{1}{2}$	0	1	0
<i>WSLS</i>	$\frac{1}{2}$	1	$\frac{(2b-1)\chi}{3b-2+(3b-1)\chi}$	$\frac{b+1}{2}$	$\frac{1-b}{2}$
$E_\chi$	0	$\frac{(2b-1)\chi}{3b-2+(3b-1)\chi}$	0	$\frac{(2b-1)\chi}{b-1+b\chi}$	0
<i>all C</i>	1	$\frac{2-b}{2}$	$\frac{2b-1}{b-1+b\chi}$	1	$1-b$
<i>all D</i>	0	$\frac{b}{2}$	0	$b$	0

where  $b$  is the benefit to the other player provided by each cooperator at the cost  $c$ , and  $\chi$  determines the surplus of the extortioner in relation to the surplus of the other player. Moreover, we use  $b - c = 1$ , thus having  $b > 1$  and  $\chi > 1$  as the two main parameters. The former determines the strength of the social dilemma, while the latter determines just how strongly strategy  $E_\chi$  exploits cooperators. A direct comparison of the extortioner strategy with the other strategies reveals that  $E_\chi$  is neutral with unconditional defectors and players adopting the *TFT* strategy. The latter, however, may beat  $E_\chi$  if they are surrounded by other *TFT* players. Similar relations hold for the competition between  $E_\chi$  and *WSLS* players. While the latter receive the same income from a direct interaction, they do gain more if the neighbors also adopt the *WSLS* strategy. It is also worth noting that the payoffs between *C* and *D* constitute the so-called donation game, which is an important special case of the iterated prisoner's dilemma game with all the original properties retained [54].

We predominantly consider a  $L \times L$  square lattice with periodic boundary conditions as the simplest interaction network to describe a structured population. To demonstrate the robustness of our findings, we also use a random regular graph and the scale-free network with the same average degree, which is likely somewhat more apt to describe realistic social and technological networks [55]. We have used pop-

ulation sizes from  $10^4$  up to  $10^6$  players to avoid finite-size effects.

Unless stated differently, for example to illustrate a specific invasion process as in Fig. 5, we use random initial conditions such that all five strategies are uniformly distributed across the network. We carry out Monte Carlo simulations comprising the following elementary steps. First, a randomly selected player  $x$  with strategy  $s_x$  acquires its payoff  $p_x$  by playing the game with its  $k$  neighbors, as specified by the underlying interaction network. Next, player  $x$  changes its strategy  $s_x$  to  $s'_x$  with the probability

$$q(s'_x \rightarrow s_x) = \frac{1}{1 + \exp[(p_x - p'_x)/K]} \quad (1)$$

where  $p'_x$  is the payoff of the same player if adopting strategy  $s'_x$  within the same neighborhood, and  $K = 0.05$  quantifies a small uncertainty that is related to the strategy adoption process [15]. The strategy  $s'_x$  should of course be different from  $s_x$ , and it is drawn randomly from the remaining four strategies. Such strategy updating is known as the myopic best response rule [51].

We also consider the more traditional strategy imitation, where player  $x$  imitates the strategy of a randomly selected neighbor  $y$ , only that  $p'_x$  in Eq. 1 is replaced by  $p_y$  [15], as well as death-birth updating as described for example in [56]. Regardless of the applied strategy updating rule, we let the system evolve towards the stationary state where the average frequency of strategies becomes time independent. We measure time in full Monte Carlo steps (*MCS*), during which each player is given a chance to change its strategy once on average.

## Acknowledgments

This research was supported by the Hungarian National Research Fund (Grant K-101490), TAMOP-4.2.2.A-11/1/KONV-2012-0051, and the Slovenian Research Agency (Grants J1-4055 and P5-0027).

[1] Maynard Smith, J. and Szathmary, E. *The Major Transitions in Evolution* (W. H. Freeman & Co, Oxford, 1995).  
[2] Maynard Smith, J. *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, U.K., 1982).  
[3] Weibull, J. W. *Evolutionary Game Theory* (MIT Press, Cambridge, MA, 1995).  
[4] Hofbauer, J. and Sigmund, K. *Evolutionary Games and Population Dynamics* (Cambridge University Press, Cambridge, U.K., 1998).  
[5] Mesterton-Gibbons, M. *An Introduction to Game-Theoretic Modelling, 2nd Edition* (American Mathematical Society, Providence, RI, 2001).  
[6] Nowak, M. A. *Evolutionary Dynamics* (Harvard University Press, Cambridge, MA, 2006).  
[7] Hamilton, W. D. Genetical evolution of social behavior I. *J.*

*Theor. Biol.* **7**, 1–16 (1964).  
[8] Nowak, M. A. and May, R. M. Evolutionary games and spatial chaos. *Nature* **359**, 826–829 (1992).  
[9] Trivers, R. L. The evolution of reciprocal altruism. *Q. Rev. Biol.* **46**, 35–57 (1971).  
[10] Axelrod, R. and Hamilton, W. D. The evolution of cooperation. *Science* **211**, 1390–1396 (1981).  
[11] Wilson, D. S. Structured demes and the evolution of group-advantageous traits. *Am. Nat.* **111**, 157–185 (1977).  
[12] Nowak, M. A. Five rules for the evolution of cooperation. *Science* **314**, 1560–1563 (2006).  
[13] Doebeli, M. and Hauert, C. Models of cooperation based on prisoner's dilemma and snowdrift game. *Ecol. Lett.* **8**, 748–766 (2005).  
[14] Sigmund, K. Punish or perish? retaliation and collaboration

- among humans. *Trends Ecol. Evol.* **22**, 593–600 (2007).
- [15] Szabó, G. and Fáth, G. Evolutionary games on graphs. *Phys. Rep.* **446**, 97–216 (2007).
- [16] Roca, C. P., Cuesta, J. A., and Sánchez, A. Evolutionary game theory: Temporal and spatial effects beyond replicator dynamics. *Phys. Life Rev.* **6**, 208–249 (2009).
- [17] Schuster, S., Kreft, J.-U., Schroeter, A., and Pfeiffer, T. Use of game-theoretical methods in biochemistry and biophysics. *J. Biol. Phys.* **34**, 1–17 (2008).
- [18] Perc, M. and Szolnoki, A. Coevolutionary games – a mini review. *BioSystems* **99**, 109–125 (2010).
- [19] Perc, M., Gómez-Gardeñes, J., Szolnoki, A., and Floría, Y. Moreno, L. M. Evolutionary dynamics of group interactions on structured populations: a review. *J. R. Soc. Interface* **10**, 20120997 (2013).
- [20] Rand, D. A. and Nowak, M. A. Human cooperation. *Trends Cogn. Sci.* **17**, 413–425 (2013).
- [21] Fudenberg, D. and Maskin, E. The folk theorem in repeated games with discounting and incomplete information. *Econometrica* **54**, 533–554 (1986).
- [22] Nowak, M. A. and Sigmund, K. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner’s dilemma game. *Nature* **364**, 56–58 (1993).
- [23] Santos, F. C. and Pacheco, J. M. Scale-free networks provide a unifying framework for the emergence of cooperation. *Phys. Rev. Lett.* **95**, 098104 (2005).
- [24] Imhof, L. A., Fudenberg, D., and Nowak, M. A. Evolutionary cycles of cooperation and defection. *Proc. Natl. Acad. Sci. USA* **102**, 10797–10800 (2005).
- [25] Santos, F. C., Pacheco, J. M., and Lenaerts, T. Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proc. Natl. Acad. Sci. USA* **103**, 3490–3494 (2006).
- [26] Tanimoto, J. Dilemma solving by coevolution of networks and strategy in a  $2 \times 2$  game. *Phys. Rev. E* **76**, 021126 (2007).
- [27] Fu, F., Liu, L.-H., and Wang, L. Evolutionary prisoner’s dilemma on heterogeneous Newman-Watts small-world network. *Eur. Phys. J. B* **56**, 367–372 (2007).
- [28] Gómez-Gardeñes, J., Campillo, M., Floría, L. M., and Moreno, Y. Dynamical organization of cooperation in complex networks. *Phys. Rev. Lett.* **98**, 108103 (2007).
- [29] Poncela, J., Gómez-Gardeñes, J., Floría, L. M., and Moreno, Y. Robustness of cooperation in the evolutionary prisoner’s dilemma on complex systems. *New J. Phys.* **9**, 184 (2007).
- [30] Fu, F., Hauert, C., Nowak, M. A., and Wang, L. Reputation-based partner choice promotes cooperation in social networks. *Phys. Rev. E* **78**, 026117 (2008).
- [31] Poncela, J., Gómez-Gardeñes, J., Floría, L. M., Moreno, Y., and Sánchez, A. Cooperative scale-free networks despite the presence of defector hubs. *EPL* **88**, 38003 (2009).
- [32] Antonioni, A. and Tomassini, M. Network fluctuations hinder cooperation in evolutionary games. *PLoS ONE* **6**, e25555 (2011).
- [33] Tanimoto, J., Brede, M., and Yamauchi, A. Network reciprocity by coexisting learning and teaching strategies. *Phys. Rev. E* **85**, 032101 (2012).
- [34] Gracia-Lázaro, C., Cuesta, J., Sánchez, A., and Moreno, Y. Human behavior in prisoner’s dilemma experiments suppresses network reciprocity. *Sci. Rep.* **2**, 325 (2012).
- [35] Gracia-Lázaro et al. Heterogeneous networks do not promote cooperation when humans play a prisoner’s dilemma. *Proc. Natl. Acad. Sci. USA* **109**, 12922–12926 (2012).
- [36] Axelrod, R. *The Evolution of Cooperation* (Basic Books, New York, 1984).
- [37] Imhof, L. A., Fudenberg, D., and Nowak, M. A. Tit-for-tat or win-stay, lose-shift? *J. Theor. Biol.* **247**, 574–580 (2007).
- [38] Press, W. and Dyson, F. Iterated prisoner’s dilemma contains strategies that dominate any evolutionary opponent. *Proc. Natl. Acad. Sci. USA* **109**, 10409–10413 (2012).
- [39] Stewart, A. J. and Plotkin, J. B. Extortion and cooperation in the prisoner’s dilemma. *Proc. Natl. Acad. Sci. USA* **109**, 10134–10135 (2012).
- [40] Adami, C. and Hintze, A. Evolutionary instability of zero-determinant strategies demonstrates that winning is not everything. *Nat. Commun.* **4**, 2193 (2013).
- [41] Hilbe, C., Nowak, M., and Sigmund, K. Evolution of extortion in iterated prisoner’s dilemma games. *Proc. Natl. Acad. Sci. USA* **110**, 6913–6918 (2013).
- [42] Stewart, A. J. and Plotkin, J. B. From extortion to generosity, evolution in the iterated prisoner’s dilemma. *Proc. Natl. Acad. Sci. USA* **110**, 15348–15353 (2013).
- [43] Szolnoki, A. and Perc, M. Reward and cooperation in the spatial public goods game. *EPL* **92**, 38003 (2010).
- [44] Helbing, D., Szolnoki, A., Perc, M., and Szabó, G. Evolutionary establishment of moral and double moral standards through spatial interactions. *PLoS Comput. Biol.* **6**, e1000758 (2010).
- [45] Szolnoki, A., Szabó, G., and Perc, M. Phase diagrams for the spatial public goods game with pool punishment. *Phys. Rev. E* **83**, 036101 (2011).
- [46] Fu, F., Tarnita, C., Christakis, N., Wang, L., Rand, D., and Nowak, M. Evolution of in-group favoritism. *Sci. Rep.* **2**, 460 (2012).
- [47] Fu, F., Nowak, M., Christakis, N., and Fowler, J. The evolution of homophily. *Sci. Rep.* **2**, 845 (2012).
- [48] Szolnoki, A. and Perc, M. Evolution of extortion in structured populations. *Phys. Rev. E* **89**, 022804 (2014).
- [49] Santos, F. C., Santos, M. D., and Pacheco, J. M. Social diversity promotes the emergence of cooperation in public goods games. *Nature* **454**, 213–216 (2008).
- [50] Hilbe, C., Nowak, M., and Traulsen, A. Adaptive dynamics of extortion and compliance. *PLoS ONE* **8**, e77886 (2013).
- [51] Matsui, A. Best response dynamics and socially stable strategies. *J. Econ. Theor.* **57**, 343–362 (1992).
- [52] Blume, L. E. The Statistical Mechanics of Strategic Interactions. *Games Econ. Behav.* **5**, 387–424 (1993).
- [53] Ellison, G. Learning, Local Interaction, and Coordination. *Econometrica* **61**, 1047–1071 (1993).
- [54] Brede, M. Short Versus Long Term Benefits and the Evolution of Cooperation in the Prisoner’s Dilemma Game. *PLoS ONE* **8**, e56016 (2013).
- [55] Barabási, A.-L. and Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
- [56] Ohtsuki, H. and Nowak, M. A. The replicator equation on graphs. *J. Theor. Biol.* **243**, 86–97 (2006).